PATENT
5681-69100
SUN030178

"EXPRESS MAIL" MAILING LABEL NUMBER
EV318247847US
DATE OF DEPOSIT SEPTEMBER 12, 2003
I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER 37 C.F.R. 1.10 ON THE DATE INDICATED ABOVE AND IS ADDRESSED TO THE COMMISSIONER FOR PATENTS, P.O. BOX 1450, ALEXANDRIA, VA 22313-1450.

Derrick Brown

A SINGLE RANK MEMORY MODULE FOR USE IN A TWO-RANK MEMORY MODULE SYSTEM

By:

Tayung Wong
Clement C. Fang
David J. Jeffrey
Neil N. Duncan

Atty. Dkt. No.: 5681-69100

Stephen J. Curran
Meyertons, Hood, Kivlin, Kowert & Goetzel, P.C.
P.O. Box 398
Austin, TX 78767-0398
Ph: (512) 853-8800

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] This invention relates to computer system memories and, more particularly, to memory module configurations.

Description of the Related Art

[0002] Many modern computer systems allow for memory expansion using memory modules. Two commonly used types of memory modules are single inline memory modules (SIMMs) and dual inline memory modules (DIMMs). SIMMs and DIMMs include small, compact circuit boards that are designed to mount easily into an expansion socket mounted on another circuit board, such as a computer system motherboard. The memory module circuit boards used to implement SIMMs and DIMMs include an edge connector that has contact pads or pins that are typically arranged on both sides of the circuit board. On SIMMs, opposing contact pads are connected together (i.e. shorted), and thus carry the same signal, while at least some of the opposing contact pads on DIMMs are not connected together, thus allowing different signals to be carried on opposing contact pads. Accordingly, DIMMs typically have a higher signal density than SIMMs.

[0003] Many memory devices mounted on SIMMs and DIMMs are devices in the Dynamic Random Access Memory (DRAM) family of devices or 'chips'. Examples of some DRAM chips include Synchronous DRAM (SDRAM) chips and Double Data Rate SDRAM (DDRSDRAM) chips. SIMMs and DIMMs are normally available in various total memory capacities. For example, they may be available in 64, 128 or 256 Megabyte capacities. The various capacities are achieved in several ways. The first is selection of

memory chips having a given address space and byte size. For example, a given chip may have a four-megabyte address space (i.e., four million separate addressable memory locations, with each location storing sixteen bits). Such a chip can provide storage of four million sixteen-bit words, and may be referred to as a 4Mx16 chip. Since memory capacity is often rated in terms of how many eight bit words the memory stores, such a chip may be considered to have eight-megabyte memory capacity. For a given size of memory chip, memory module capacity can be increased by using multiple chips on a board and increasing data bus width so that the data at the same addressed location in each chip can be read out to the bus simultaneously. For example, if three 4Mx16 chips are used, the bus width would need to be at least forty-eight to allow all of the bits at a selected address to be read out to the bus at the same time. A module with three 4Mx16 chips can be considered to have a total capacity of twelve million sixteen-bit bytes, but may be called a twenty-four megabyte memory in terms of eight-bit bytes.

[0004] From the above examples, it may be shown that reliability of a memory module may decrease with an increase in memory capacity and data path width. If the data path width increases due to an increase in the number of memory chips used to create the width, then the failure rate for the memory module increases. One method of calculating the average life of the memory module is the parts count method. Using the parts count method, the mean time between failure (MTBF) of an assembly (e.g., memory module) is calculated using the inverse of the sum the individual failure rates of each component. Thus, if the number of components increases, then the failure rate for the memory module increases and the MTBF decreases.

[0005] In addition, as higher capacity memory devices become available in the market, at some point in the higher capacity memory device timeline, the higher capacity devices enter what is referred to as the "sweet spot" of the price/performance charts. This generally means to the purchaser, that the device cost on a per bit basis gives the best cost

for a given amount of performance. Thus, if a purchaser continues to use older smaller capacity devices, the will move out of the sweet spot since the performance may no longer be in the desired range for a given cost model. Thus, it is generally desirable to use devices that are in the sweet spot of the price/performance chart. To that end, it may make sense to transition to using higher capacity memory devices on a given memory module even while the given memory module is still widely used. However in many cases, doing so may present system compatibility issues.

SUMMARY OF THE INVENTION

[0006] Various embodiments of a memory module for use in a two rank memory module system are disclosed. In one embodiment, the memory module includes a plurality of memory devices and a control circuit. The control circuit may be configured to generate a chip select signal that is provided to each of the memory devices. The chip select signal may be dependent upon assertions of a first bank chip select signal and a second bank chip select signal received from a memory controller. The control circuit may be further configured to generate an address signal that is provided to each of the memory devices. The address signal may be asserted dependent upon which of the first bank chip select signal and the second bank chip select signal are asserted.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] FIG. 1 is a block diagram of one embodiment of a computer system.

[0008] FIG. 2 is a diagram illustrating one embodiment of a two-rank memory module.

[0009] FIG. 3 is a diagram illustrating one embodiment of a single rank memory module with the same storage capacity of the memory module of FIG. 2.

[0010] FIG. 4 is a pin assignment table for the memory modules of FIG. 2 and FIG. 3.

[0011] FIG. 5A is a diagram illustrating one embodiment of the buffer circuit of FIG. 3.

[0012] FIG. 5B is a diagram illustrating one embodiment of the control circuit of FIG. 3.

[0013] While the invention is susceptible to various modifications and alternative forms, specific embodiments thereof are shown by way of example in the drawings and will herein be described in detail. It should be understood, however, that the drawings and detailed description thereto are not intended to limit the invention to the particular form disclosed, but on the contrary, the intention is to cover all modifications, equivalents and alternatives falling within the spirit and scope of the present invention as defined by the appended claims.

DETAILED DESCRIPTION

[0014] Turning now to FIG. 1, a block diagram of one embodiment of a computer system 10 is shown. Computer system 10 includes a plurality of processors 20-20n connected to a memory subsystem 50 via a system bus 25. Memory subsystem 50 includes a memory controller 30 coupled to a system memory 40 via a memory bus 35. It is noted that, although two processors and one memory subsystem are shown in FIG. 1, embodiments of computer system 10 employing any number of processors and memory subsystems are contemplated. In addition, elements referred to herein with a particular reference number followed by a letter may be collectively referred to by the reference number alone. For example, processor 20A-n may be collectively referred to as processor 20.

[0015] Memory subsystem 30 is configured to store data and instruction code within system memory 40 for use by processor 20. As will be described further below, in one embodiment, system memory 40 may be implemented using a plurality of memory modules 100. In one embodiment, memory modules 100 may be implemented using DIMMs. As noted above, each DIMM may employ a plurality of memory chips in the DRAM family of chips as SDRAM or DDRSDRAM chips, for example. Each DIMM may be mated to a system memory board via an edge connector and socket arrangement. In one embodiment, the socket may be located on a memory subsystem circuit board and each DIMM may have an edge connector which may be inserted into the socket, for example.

[0016] Generally speaking, processor 20 or an I/O device (not shown) may access memory subsystem 50 by initiating a memory request transaction such as a memory read or a memory write to memory controller 30 via system bus 25. Memory controller 30 may then control the storing to and retrieval of data from system memory 40 by issuing

memory request commands to system memory 40 via memory bus 35. Memory bus 35 conveys address and control information and data between system memory 40 and memory controller 30. The address and control information may be conveyed to each DIMM in a point-to-multipoint arrangement while the data may be conveyed directly between each memory chip on each DIMM and memory controller 30 in a point-to-point arrangement. The point-to-multipoint arrangement is sometimes referred to as a multi-drop topology.

[0017] In one embodiment, memory bus 35 includes address and control information such as address signals AD0 – AD15, column address strobes CAS0 and CAS1, row address strobes RAS0 and RAS1, write enable signals WE0 and WE1, bank chip select signals CS0 and CS1 and clock signals CLK+ and CLK-. In addition in one embodiment, memory bus 35 may include a 576-bit data path including data signals DQ0-DQ575, with 144 (e.g., DQ0-143 to DIMM 0, DQ144-287 to DIMM 1, and so on) data pins being routed between memory controller 30 and each DIMM. It is noted that depending on the implementation, address signals AD12-AD15 received on the DIMM connector from memory controller 30 may be used as bank select signals BS0 and BS1, and address signals AD12 and AD13 on the memory devices.

[0018] Generally speaking, SDRAM devices may be accessed in a burst-oriented mode. Typically, accesses start at a selected location and continue for a programmed number of locations in a programmed sequence. Accesses begin with the registration of an ACTIVE command, which is then followed by a READ or WRITE command. The address bits registered coincident with the ACTIVE command are used to select the bank and row to be accessed. For example, BA0 and BA1 may select the bank, and A0- A11 may select the row. The address bits registered coincident with the READ or WRITE command are used to select the starting column location for the burst access. A command typically includes an encoding of signals such as the CS, RAS, CAS and WE

along with the appropriate address and data signals. Prior to normal operation, most SDRAMs must be initialized, which includes having their respective mode register programmed.

[0019] As described above, memory devices come in a variety of sizes and configurations. Further, it may be desirable to transition to higher capacity memory devices on a given memory module and to continue to use that same memory module in a given system. However as noted above, doing so may introduce compatibility issues. One example of such a compatibility issue is as follows: If a given memory module has a 256 MB capacity using a number of 128MB SDRAM devices in an 8M x 16 configuration, reconfiguring the module to use fewer, but higher capacity memory devices (e.g., 256MB SDRAM devices in a 16M x 16 configuration) while still providing an overall 256MB capacity may change the entire address and control scheme for the memory module. Thus, in computer system 10, memory subsystem 30 may be incapable of using the new memory module. To avoid such compatibility issues, the memory modules having the higher capacity memory devices must be made to appear the same as the memory modules using the lower capacity memory devices to memory controller 30.

[0020] Most memory modules may come in either a single rank or a dual rank configuration. The rank typically refers to whether both sides of the memory module are populated with memory chips and how the memory chips are logically arranged. For example, a single rank module may only have memory chips located on one side and a dual rank module may have memory chips located on both sides of the module. Further, on some dual rank modules, each rank may be accessed via a different address and control signals (e.g., a different chip select signal). Thus, a 256MB capacity memory module may be implemented using two ranks of 128MB memory chips. Alternatively, a 256MB capacity memory module may be implemented using one rank of 256MB memory chips. However in conventional computer systems, a single rank memory

module may not be used in a system designed for a dual rank memory module and vice versa.

[0021] As will be described in greater detail below in conjunction with the descriptions of FIG. 2 and FIG. 3, the memory module of FIG. 2 is implemented as a 256MB module using 128MB SDRAM devices in a dual rank 8MB x16 configuration and the memory module of FIG. 3 is implemented as a 256MB module that is compatible with the memory module of FIG. 2 while using higher capacity 256MB SDRAM devices in a single rank 16MB x16 configuration. It is noted that in one embodiment, the memory module of FIG. 2 and the memory module of FIG. 3 may be compatible with a family of memory modules that are referred to as next generation (NG) DIMMs by Sun Microsystems, Inc.

[0022] Referring to FIG. 2, a block diagram of one embodiment of a two-rank memory module is shown. Memory module 100A includes a circuit board with eighteen memory devices 210 arranged into two banks (e.g., bank 0 and bank 1) mounted to it. In addition, memory module 100A includes a buffer circuit 220 and a clock circuit 250 that are coupled to the memory devices 210. Further, memory module 100A includes a serial presence detect (SPD) device 275. It is noted that in one embodiment, each bank may be physically located on a given side of the circuit board of memory module 100A.

[0023] In the illustrated embodiment, the nine memory devices 210 included within bank 0 are coupled to the data path (e.g., DQ0-DQ143). In addition, the nine memory devices 210 included within bank 1 are also coupled to the data path (e.g., DQ0-DQ143). Specifically, a given memory device 210 in each bank is coupled to a same respective set of data signals (e.g., DQ0-15). Thus, every set of data signals has two memory devices 210 (one from each bank) coupled to it.

[0024] In the illustrated embodiment, memory module 100A is configured to receive address, control and data signals as well as a serial data from a memory controller such as memory controller 30 of FIG. 1, for example. In the illustrated embodiment, memory module 100A is coupled to address signals A0-A13 and control signals CAS0/, CAS1/, RAS0/, RAS1/, CKE0, CKE1, WE0, WE1, CS0/, CS1/, CLK+, and CLK-. In addition, memory module 100A is coupled to 144 data signals. In the illustrated embodiment, the data signals are designated as DQ0-DQ143. However, as described above, each memory module may include 144 data signals and they may be designated DQ144-287, or DQ288-431, etc... Memory module 100A is also coupled to a serial interface such as an Inter-IC (I2C®) bus, developed by Philips Semiconductors, Inc., for example, for connection between SPD 275 and the memory controller. It is noted that signal names followed by a '/' are indicative of active low signals.

[0025] Buffer circuit 220 is used to buffer the address and control signals received at the DIMM socket from the memory controller and to drive the address and control signals to memory devices 210. In one embodiment, to adequately drive memory devices 210, each input to buffer circuit 220 may have two or more duplicate corresponding outputs. However, other embodiments may have a single output for each input.

[0026] In one embodiment, the memory devices 210 associated with each bank (e.g., bank 0 and bank 1) may be accessed using the bank chip select signals CS0/ and CS1/, respectively. Thus, each bank may be associated with a given address range. Accordingly, one bank may be accessed using one bank chip select on one clock edge and the other bank may be accessed using the other bank chip select on a next clock edge. This may allow data to be retrieved every two clocks.

[0027] During operation, the memory controller sends complete sets of control signals for each bank (e.g., CAS0/ and CAS1/). In addition, since two banks are used, only

address signals A0-A11 are necessary to address 256 MB of addressable space. One bank may be used to store the lower 128MB and the second bank may be used to store the upper 128MB of addressable space with the chip selects being used to select between the two. Address signals AD12 and AD13 are used as memory device internal bank select signals BS0 and BS1.

[0028] Generally speaking, the DRAM family of devices such as memory devices 210 must be refreshed periodically at a refresh rate that is generally manufacturer specific. SDRAM-type devices typically decode a refresh command sent by the memory controller. The refresh command may also be manufacturer specific, but it generally includes decoding one or more of the control signals in combination with an active chip select into a refresh command. In one embodiment, the bank 0 memory devices 210 may be refreshed by decoding a refresh command from the memory controller in combination with an active CS0/ signal, while the bank 1 memory devices 210 may be refreshed by decoding a refresh command from the memory controller in combination with an active CS1/ signal.

[0029] As described above, memory module 100A may be implemented as a two rank 256MB module using two banks of nine 128MB memory devices. Thus, each of memory devices 210 may be a 128MB SDRAM device in a 8MB x16 configuration. In one embodiment, the data path of 144 data signals includes 128 data signals and 16 redundant data signals. Thus, for each bank of nine memory devices 210, one memory device may be used to store redundant data.

[0030] Clock circuit 250 may be configured to generate clock signals derived from the differential clock signals CLK+ and CLK-. For example, in one embodiment, clock circuit 250 may include a phase locked loop (PLL) circuit configured to generate clock signals for memory devices 210.

[0031] SPD 275 may be configured to provide information from memory module 100A to the memory controller corresponding to the memory module's properties and functionality. For example, SPD 275 is a storage unit that may store information such as the number of row addresses and column addresses for memory devices 210, the number of banks, etc. The ability to read this information may allow the memory controller to be configured specifically for the type of memory module being used. In one embodiment, SPD 275 is a serially accessible storage such as an electrically erasable programmable read only (EEPROM) device.

[0032] Turning to FIG. 3, a block diagram of one embodiment of a single rank memory module is shown. Memory module 100B includes a circuit board having nine memory devices 310 arranged into one bank (e.g., bank 0) mounted to it. In addition, memory module 100B includes a buffer circuit 320, a clock circuit 350 and a control circuit 330 that are coupled to the memory devices 310. Further, memory module 100B includes SPD device 375.

[0033] In the illustrated embodiment, there are nine memory devices 310 included within bank 0 and they are coupled to the data path (e.g., DQ0-DQ143). Specifically, a given memory device 310 is coupled to a respective set of data signals (e.g., DQ0-15). Thus, every set of data signals has one memory device 310 coupled to it.

[0034] In the illustrated embodiment, memory module 100B is configured to receive address, control and data signals as well as a serial data from a memory controller such as memory controller 30 of FIG. 1, for example. In the illustrated embodiment, memory module 100B is coupled to address signals A0-A13 and control signals CAS0/, CAS1/, RAS0/, RAS1/, CKE0, CKE1, WE0, WE1, CS0/, CS1/, CLK+, and CLK-. In addition, memory module 100B is coupled to 144 data signals. In the illustrated embodiment, the

data signals are designated as DQ0-DQ143. However, as described above, each memory module may include 144 data signals and they may be designated DQ144-287, or DQ288-431, etc. Memory module 100B is also coupled to a serial interface such as an Inter-IC (I2C®) bus, developed by Philips Semiconductors, Inc., for example, for connection between SPD 375 and the memory controller.

[0035] As will be described in greater detail below in conjunction with the description of FIG. 5A, buffer circuit 320 is used to buffer the address and control signals received at the DIMM socket from the memory controller and to drive the address and control signals to memory devices 310. In one embodiment, to adequately drive memory devices 310, each input to buffer circuit 320 may have two or more duplicate corresponding outputs. However, other embodiments may have a single output for each input.

[0036] Since memory module 100B is compatible with memory module 100A, during operation, the memory controller sends complete sets of control signals for each bank (e.g., CAS0/ and CAS1/, etc...). However, since only one bank is used on memory module 100B, address signals A0-A11 are not sufficient to address the full 256 MB of addressable space. Further, since address signals AD12 and AD13 are used as memory device internal bank select signals BS0 and BS1, an AD12 signal must be generated on memory module 100B.

[0037] As will be described in greater detail below in conjunction with the description of FIG. 5B, control circuit 330 includes logic which may be configured to decode the bank chip select signals CS0/ and CS1/ and to generate a single chip select signal designated CS/ for enabling the single bank of memory devices 310. In addition, control circuit 330 may be configured to generate an address signal designated AD12 for addressing the upper 128MB of addressable locations. In one embodiment, control circuit 330 may be implemented as a programmable logic device (PLD) such as an

erasable PLD (EPLD), although other embodiments are contemplated in which control circuit 330 is implemented as either one or more dedicated application specific integrated circuits (ASIC) or discrete combinatorial logic devices.

[0038] As described above in conjunction with the description of FIG. 2, memory devices 210 must be refreshed periodically at a refresh rate that is generally manufacturer specific. Similarly, memory devices 310 of FIG. 3 must also be refreshed periodically. In one embodiment, memory devices 310 may be refreshed by control circuit 330 decoding a refresh command from the memory controller in combination with an active CS0/ signal and generating an appropriate CS/ signal. Since the memory controller may be configured to refresh two banks, control circuit 330 may be configured to ignore the refresh command sent for the nonexistent second bank.

[0039] As described above, memory module 100B may be implemented as a single rank 256MB module using one banks of nine 256MB memory devices. Thus, each of memory devices 310 may be a 256MB SDRAM device in a 16MB x16 configuration. In one embodiment, the data path of 144 data signals includes 128 data signals and 16 redundant data signals. Thus, for the bank of nine memory devices 310, one memory device may be used to store redundant data.

[0040] Clock circuit 350 may be configured to generate clock signals derived from the differential clock signals CLK+ and CLK-. For example, in one embodiment, clock circuit 350 may include a phase locked loop (PLL) circuit configured to generate clock signals for memory devices 310.

[0041] SPD 375 may be configured to provide information from memory module 100A to the memory controller corresponding to the memory module's properties and functionality. For example, SPD 375 is a storage unit that may store information such as

the number of row addresses and column addresses for memory devices 310, the number of banks, etc. In addition, SPD 375 may store the rank, storage capacity and error history of memory module 100B. The ability to read this information may allow the memory controller to be configured specifically for the type of memory module being used and to aid the system processor or service processor (not shown) in determining the cause of memory errors. In one embodiment, SPD 375 is a serially accessible storage such as an electrically erasable programmable read only (EEPROM) device.

[0042] It is noted that, by using a single rank module having the same capacity as a two-rank module but using fewer memory devices, a savings in power and cost may be realized. In addition, a reduced module failure rate, which translates to an increased MTBF may be realized. Further, the use of either module may be transparent to the user, since the single rank module may 'look' the same as the two-rank module to the memory controller due to the use of control circuit 330.

[0043] Turning now to FIG. 4, a pin assignment table for the memory modules of FIG. 2 and FIG. 3 is shown. As described above, memory modules 100A and 100B may both be compatible with the family of DIMMs referred to as NG DIMMs. Accordingly, both memory modules 100A and 100B may have edge connectors having the same contact pin assignment. Memory module pin assignment table 400 is an exemplary diagram illustrating the numbering of the contact pins and the signals assigned to the corresponding pin numbers associated with the edge connector of memory module 100A and 100B of FIG. 2 and FIG. 3, respectively. Pin assignment table 400 includes multiple rows and columns. The columns are labeled Pin Number and Pin Name. Therefore, each pin number has a corresponding pin name associated with it. For example, Pin number 1 is referenced as VSS. Pin number 2 is referenced DQ0 and so forth. In this particular embodiment, the ground pins are referenced VSS and the power pins are referenced

VDD. Further, the data signal pin names are DQ0-DQ143 and the address signal pin names are AD0-AD15.

[0044] Referring to FIG. 5A, a diagram of one embodiment of the buffer circuit of FIG. 3 is shown. Buffer circuit 320 includes one or more integrated circuits including a plurality of inputs and outputs configured to buffer received address and control signals and to drive the address and control signals.

[0045] In one embodiment, buffer circuit 320 is a BiCMOS line driver IC. Each input may drive two outputs, thereby decreasing the load on each output. Accordingly, each output may drive either four or five memory devices 310.

[0046] Referring to FIG. 5B, a diagram of one embodiment of the control circuit of FIG. 3 is shown. Control circuit 330 includes combinatorial logic configured to decode the received chip select signals CS0/ and CS1/ into a single chip select CS/ and an AD12 signal. For example, an exemplary truth table for the chip select and address decode is shown in Table 1 below.

Table 1: Chip Select truth table

CS0/	CS1/	CS/	AD12
L	L	L	H
L	H	L	L
H	L	L	H
H	H	H	L

[0047] According to the truth table shown above, in response to receiving an active CS0/ signal, the lower 128MB addresses of memory devices 310 are accessible via an

active CS/ signal generated by control circuit 330. In addition, in response to receiving an active CS1/, the upper 128MB addresses of memory devices 310 are accessible via an active CS/ signal and an active AD12 signal generated by control circuit 330.

[0048] It is noted that in other embodiments, control circuit 330 may include sequential elements such as flip-flops (not shown), for example, which may be used to latch or capture received signals.

[0049] In addition, in the illustrated embodiment, control circuit 330 includes refresh enable logic 331 that is configured to decode the received control signals into a refresh command for bank 0 only. A refresh command targeted for bank 1 is ignored. Accordingly, in response to receiving control signals that are decoded into a refresh command, a CS/ is generated by refresh enable circuit 331.

[0050] It is noted that the AND gate A1 and inverter I1 are representative of the combinatorial logic included within control circuit 330 and that other embodiments may include other logic circuit arrangements having similar functionality. Accordingly, in one embodiment, an active CS/ signal may be generated by control circuit 330 in response to receiving the combination of CS0/ and CS1/ signals shown in the truth table above and in response to decoding a refresh command in combination with an active CS0/ (i.e. refresh to bank 0).

[0051] Although the embodiments above have been described in considerable detail, numerous variations and modifications will become apparent to those skilled in the art once the above disclosure is fully appreciated. It is intended that the following claims be interpreted to embrace all such variations and modifications.